Jet Propulsion Laboratory
California Institute of Technology

IT Chief Technology and Innovation Office (org 176)

NASA JPL: Small Satellite Data Science Pilot

Advanced IT Research and Open Source Projects Office (1761)

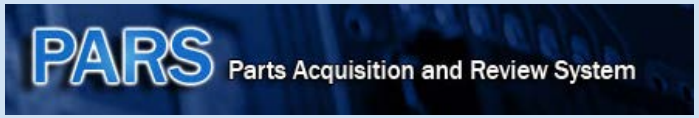
NASA NEPP ETA 2018 Workshop

NASA GSFC

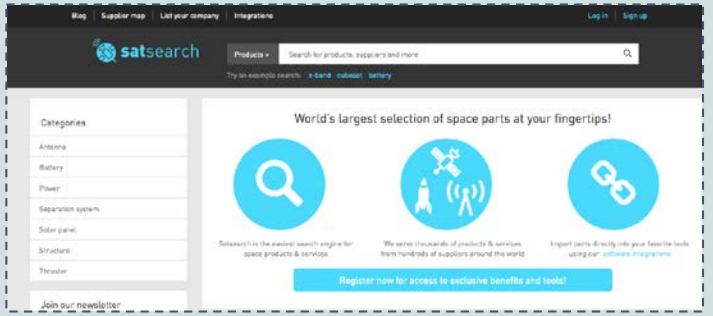
June 2018

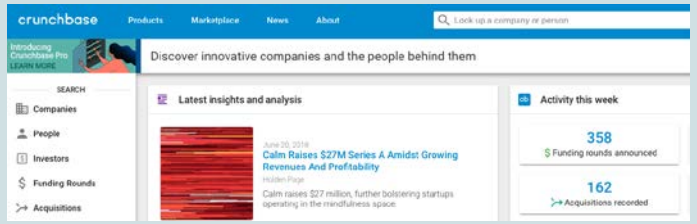
Information about small satellites is distributed across disparate resources

NASA



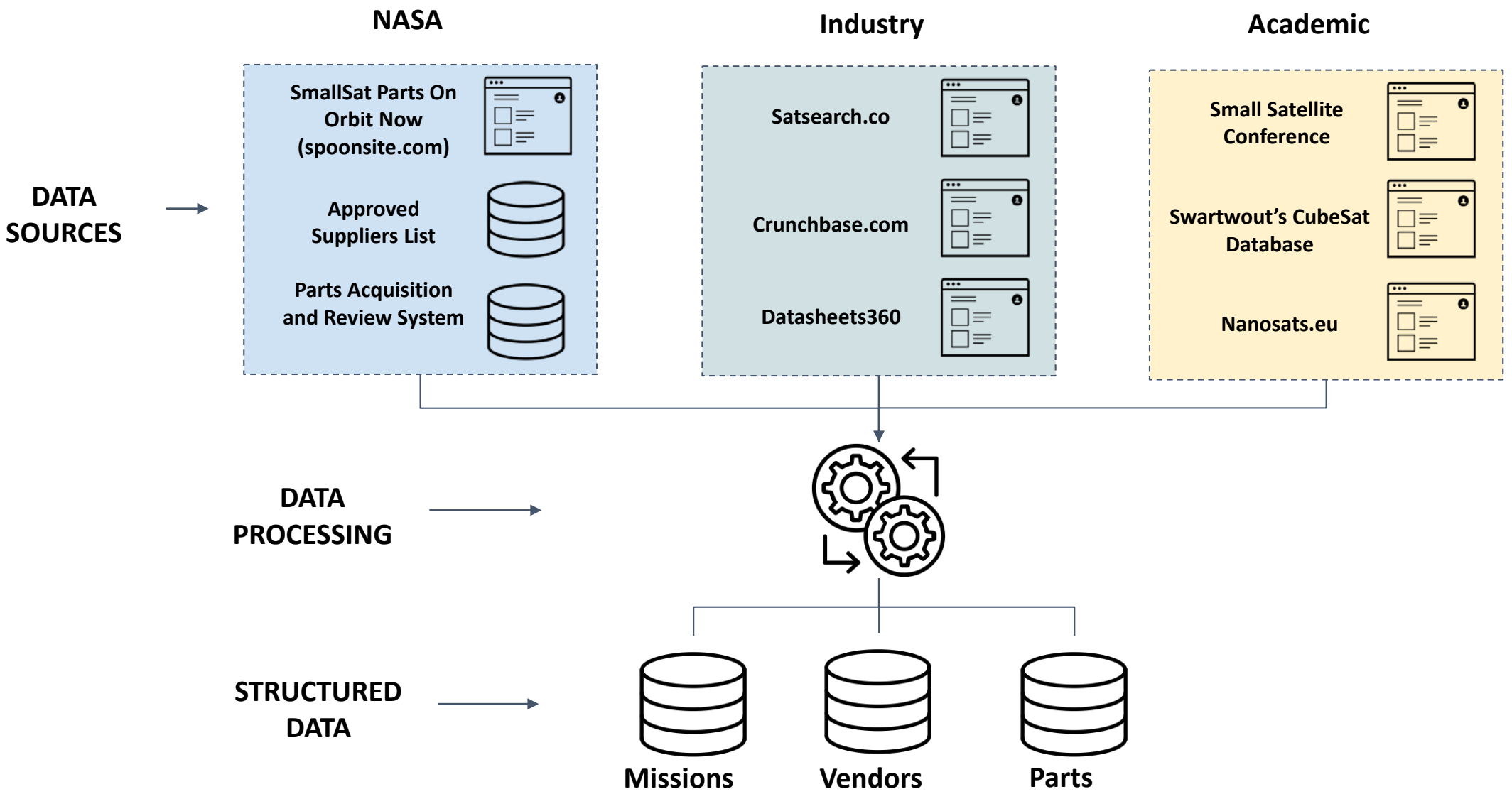
Industry





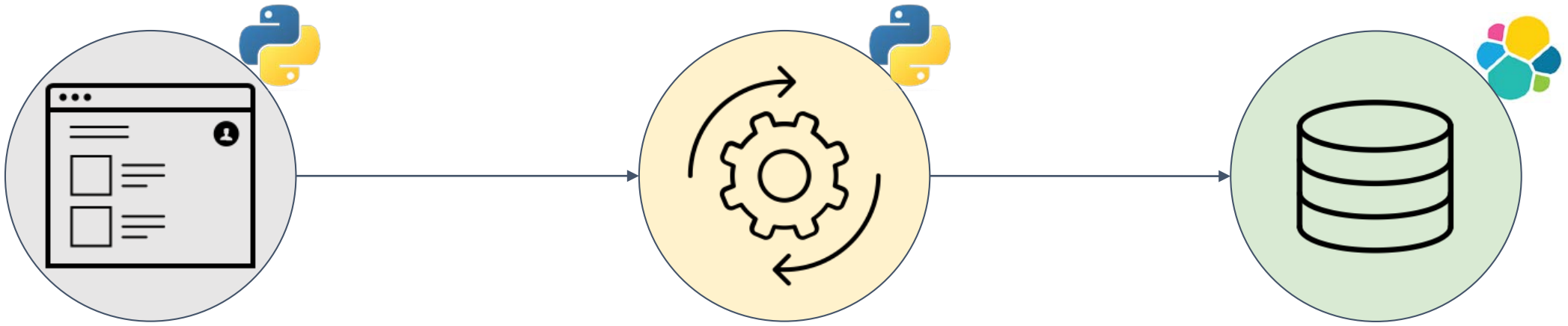
Academic

Data Collection Process Overview



Web Scraping Content

Cubesats.org, Swartout Cubesat DB, etc.



Web Scraping to Automate Data Collection Across Disparate Data Sources

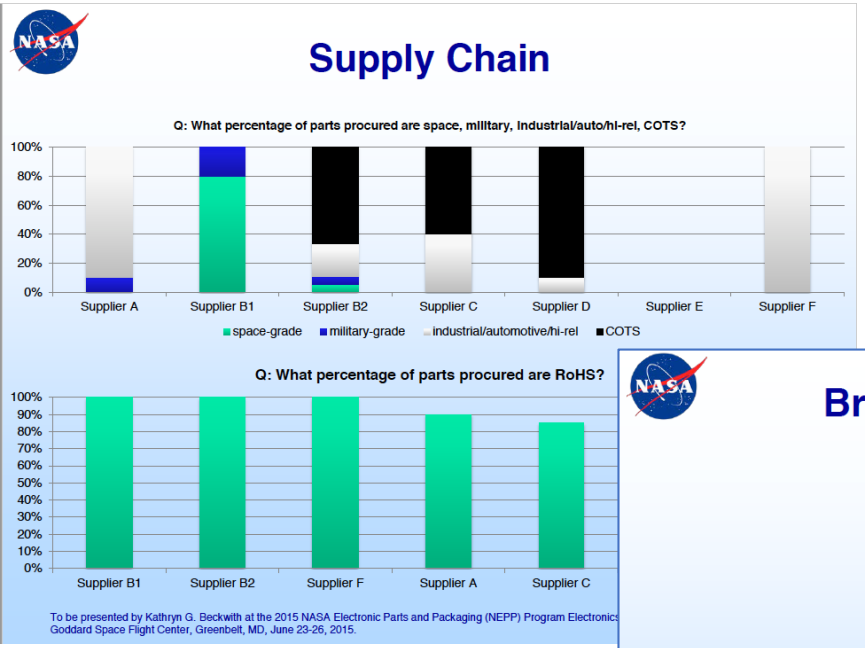
Processing and Cleansing Structured and Unstructured Data

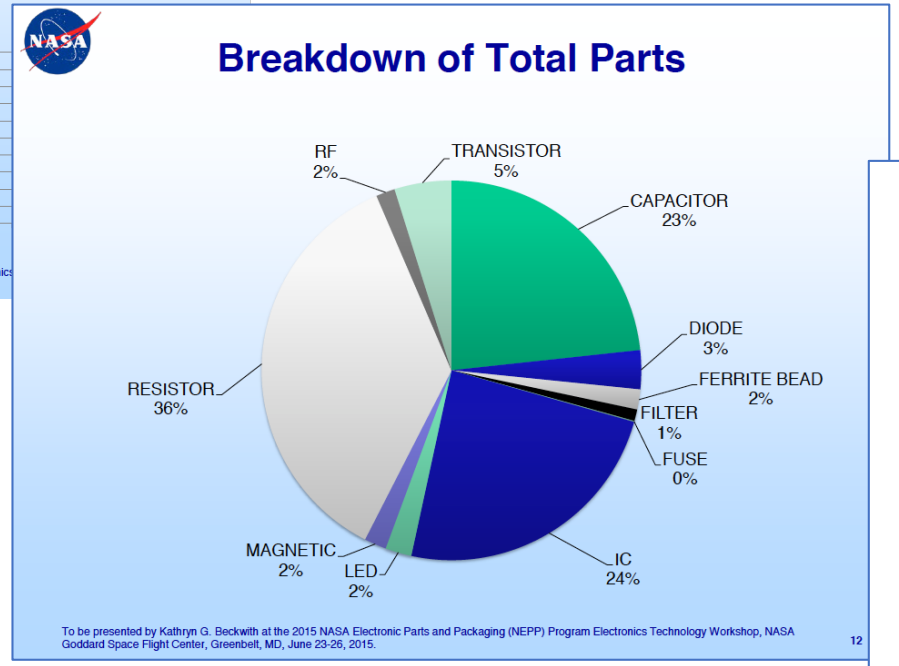
Indexing Clean Data in Elastic

NEPP SmallSat Supplier Task History

- **2015-2016 (Beckwith/Smith)**
  - Survey of 5 SmallSat Suppliers (integrators and product providers)
    - 5 Questions
    - Quality criteria based on ISO9001 standards
  - Database of NASA and JPL EEE parts usage on smallsats
- **2016-2017 (Sundgaard)**
  - Continued and expanded:
    - EEE Parts database for usage on NASA Smallsat missions
    - Survey of suppliers
      - 12 vendors, quantitative rankings
- **2018 (Mattmann et. al.)**
  - Formal data science technqiues & methodologies applied to broad range of data sources

Results of historical Small Sat Supplier Task





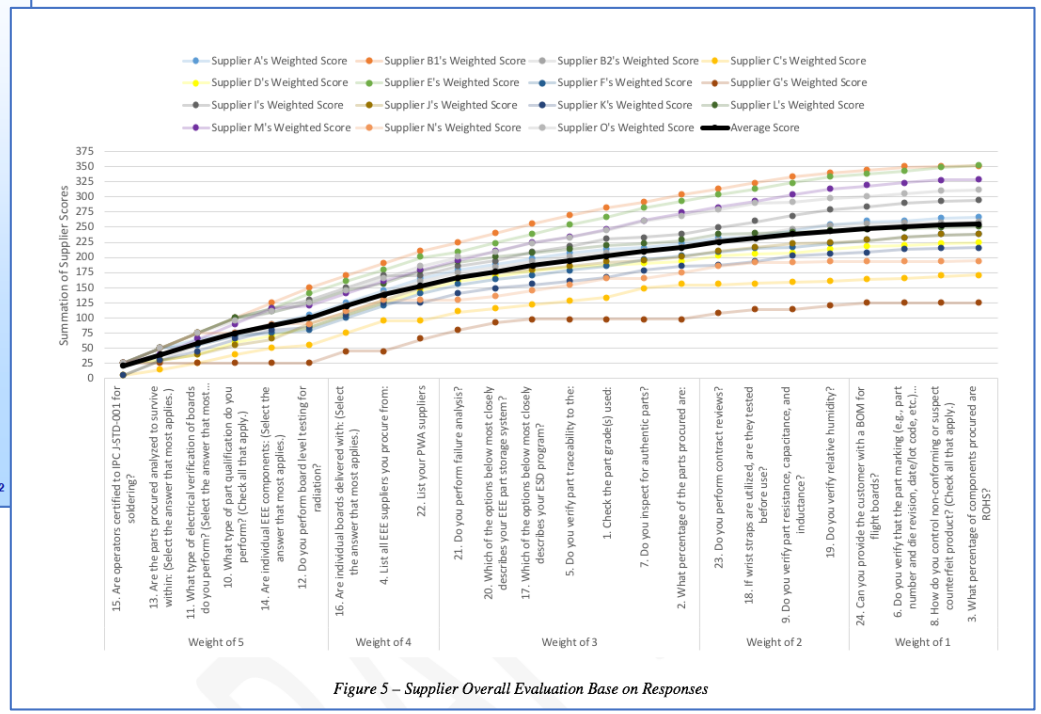
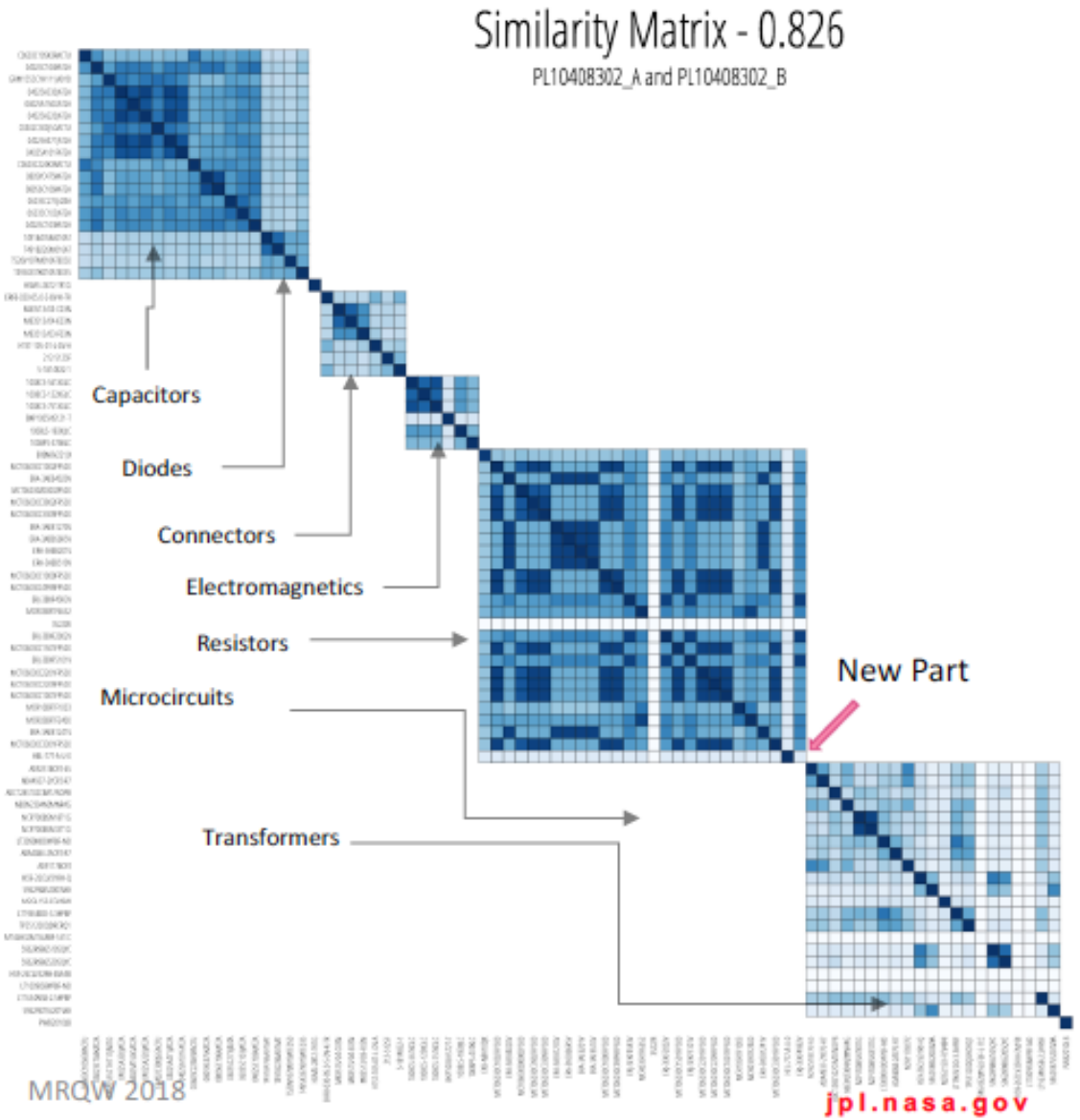


Figure 5 – Supplier Overall Evaluation Base on Responses

Cosine Similarity – Parts List Revision

- **Similarity matrix for two part lists**
- **Coloring indicates strength of similarity:**
  - 0 – white (no similarity)
  - 1 – dark blue (identical)
- **Diagonal of 1's is an artifact of comparing different revisions of the same part list**
  - One-off in diagonal indicates addition of a **new part**
- **Distinct regions correspond to part types**
  - Part comparisons are not made between different part types
- **Manage subtle changes in part type, provide an precise definition of "Heritage" as way to reduce risk**



End Game: Pilot

- Situational Awareness of Vendor / Startup space in Small Sats related to parts used in Missions
- Better exploratory metrics
  - Cosine similarity, but also other feature similarities
    - Explore Jaccard, Edit Distance, etc.
  - Clustering techniques, similar parts, vendors, and relationships
  - Ranking algorithms for exploring vendor space and parts
- Ultimate Goal: better understanding of supply chain as it relates to our missions

Thanks!

JPL Small Sat Data Science Team